

Sequence Listing

Individual Applicant

Street :
 City : Athens
 State : GA
 Country : USA
 PostalCode : 30602-7272
 PhoneNumber : 706-542-0928
 FaxNumber : ____-____-____
 EmailAddress : wparrott@uga.edu
 <110> LastName : Parrott
 <110> FirstName : Wayne
 <110> MiddleInitial : A
 <110> Suffix : PhD.

Individual Applicant

Street :
 City : Athens
 State : GA
 Country : USA
 PostalCode : 30602-7272
 PhoneNumber : 706-542-0928
 FaxNumber : ____-____-____
 EmailAddress :
 <110> LastName : LaFayette
 <110> FirstName : Peter
 <110> MiddleInitial : R
 <110> Suffix : PhD.

Individual Applicant

Street : 1907 South Milledge Ave
 City : Athens
 State : GA
 Country : USA
 PostalCode : 30605-____
 PhoneNumber : ____-____-____
 FaxNumber : ____-____-____
 EmailAddress :
 <110> LastName : Kane
 <110> FirstName : Patrick
 <110> MiddleInitial : M
 <110> Suffix :

Application Project

<120> Title : Arabitol or Ribitol As Positive Selectable Markers
 <130> AppFileReference : UGA-855R
 <140> CurrentAppNumber :
 <141> CurrentFilingDate : ____-__-__

Sequence

<213> OrganismName : Escherichia coli
 <400> PreSequenceString :
 atgaacgaac aatttacatg gctgcacatc gggtaggtt cttttcatcg cgcacatcag 60
 gcgtgggtatc tacaccgttt gcaggtgatg ggcgataaac gctggagcat tgctgcgggc 120
 aatattcgta atgatgctga acatgtcgtg caggcactca gtgcacagaa aggtcgctat 180
 gtgctggaaa ccgtcagccc ggaaggggta agcgaatatg aagagatcac ctcaattcag 240

aagttgatac	cgtggcaggc	agatttacaa	ccgctgattg	ctgaaggggc	agatccgaag	300
acaaaagtga	ttgctttcac	cgtcaccgaa	ggcgggtact	acctgaatac	cagtcacaaa	360
ctggaagtta	acaatcctga	tttagcggca	gatcttaaag	ggggatgcaa	aacaatttac	420
ggtgttatta	cccgtatcct	cgaagcgcgt	atggcaaata	acgccggacc	actaaccctg	480
atgaacgaac	aatttacatg	gctgcacatc	gggttaggtt	cttttcacatg	cgcacatcag	540
gcgtggatc	tacaccgttt	gcaggtgatg	ggcgataaac	gctggagcat	tgctcggggc	600
aatatctcga	atgatgtcga	acatgtcgtc	caggcaactc	gtcgacagaa	aggtcgctat	660
gtgctgga	ccgtcagccc	ggaaggggta	agcgaatatg	aagagatcac	ctcaattcag	720
aagttgatac	cgtggcaggc	agatttacaa	ccgctgattg	ctgaaggggc	agatccgaag	780
acaaaagtga	ttgctttcac	cgtcaccgaa	ggcgggtact	acctgaatac	cagtcacaaa	840
ctggaagtta	acaatcctga	tttagcggca	gatcttaaag	ggggatgcaa	aacaatttac	900
ggtgttatta	cccgtatcct	cgaagcgcgt	atggcaaata	acgccggacc	actaaccctg	960
ctgaagtgcg	ataacgtgcg	ccataatggt	gaacgtttcc	atgatggcct	ggttgagttt	1020
ctccagctaa	ctggcacaac	ggatgtcatc	gactggctga	gtacaaatac	cacttgcccc	1080
aataccatgg	ttgaccgcat	tacgcctcgt	ccggcagcag	aacttcgggc	acggatcaag	1140
gctcaaacgg	gtattgccga	taaagcgccg	gtaatgggcg	aaacctttat	ccagtgggtc	1200
gtggaagata	atttcctgta	tgtccgtccg	gcactggaga	aggtcgggtg	cgaactgggtg	1260
gcgtcggtaa	tcccctatga	agaggcgaag	attcgcattc	ttactcttc	acacagttgc	1320
atcgctggg	caggtacgtt	aatcggtc	aaatatatcc	acgaaagcac	aatgaccgat	1380
ttctctatc	agattgccga	cgcgtactg	gtgcaagatg	tcattccttg	cttgggcgat	1440
aatacgtatc	atttggcca	ctaccgtgat	ctgttaactc	agcgttttac	caatccacat	1500
attcaggaca	ccaaccaacg	cgtcgctgcg	gatggtttct	cgaaaattcc	ggcgaatgatt	1560
gccccacac	tgcgagagt	ctaccagcga	ggcgttcgcc	cgaatgccac	cgccatgtta	1620
cctgcactgt	tttacgtatt	catggagcag	tggcatcacg	gcaaactgcc	ctatgaatat	1680
caggatggca	tccttgatgc	accagctgtc	catgcaatgt	tacagtctgc	cgatcccgtc	1740
gctgtttatg	ccagtgataa	agcgctgttt	ggcgatttaa	ccgaactgta	agattttgcc	1800
gcgtgtttgc	gcgaaaaaat	cctgcacgtc	tacgcgttaa	ttacttaa		1848

<212> Type : DNA

```
<211> Length : 1848
```

SequenceName : SEQ ID NO 1

SequenceDescription : arabitol dehydrogenase

Custom Codon

Sequence Name : SEQ ID NO 1

Sequence

<213> OrganismName : Escherichia coli

```
<400> PreSequenceString :
```

atcgattgag	cagtttgcct	cacacggcag	ctaaattccc	gttcagtgcg	tgcaagcaaa	60
cagagactat	aaattcgcgc	tggtaaaagg	attatatgat	gaatcactct	gtgccctcta	120
tgaatactcc	ccttaatggc	aaagttgcag	ctatcactgg	cgctgcgtca	ggtattggcc	180
tgcaattgtc	aaaaacgctg	ctcgcgtcag	gagcaaaagg	agtattgatt	gaccgggaag	240
gcgcaaaact	gcacaagatt	gtcgtcgtag	taggcgaaaa	cgcgtacgcg	ctgcaactcg	300
atctcttcaa	taatcagcaa	gtcgataaca	tgttcggcga	cattatcgaa	ctggcggtg	360
ggctggatat	ttttcatgcc	aatgcaggcg	cttatattgg	cggcccagtg	gctgaagggtg	420
atccagatgt	ctgggatcgt	gtgttaaata	tgaatataaa	tgcggcggtt	cgctgtgtcc	480
gtgcagtcct	gccgcataat	attgcgcaga	ggtcgggcga	tataattttt	accagttcca	540
tcgcggggcg	cgttccgggt	atctgggaac	cgatctacac	cgcgtccaaa	tttgccggtc	600
aggcattcgt	acacactacc	cgcgcgccag	tttctcaata	tggcgtgcgt	gtgggtgcgg	660
tgctgccagg	accagtagtc	actgcctcgt	ttgatgactg	gccaaaagcc	aaaaaggaa	720
aagccctggc	aaatggtagc	ctgatgcaac	cgattgaagt	ggcggaaatc	gtattgttta	780
tggtgaccgc	ctcgaaaaat	gtcacctgtc	gagatttagt	gatcctgcct	ggcagtgctg	840
atctgtaagg	gcgcaatcat	gacaataacc	aaaacggtta	ttggtgttga	tgtgggatca	900
ggcagtgctc	gcgcgcggat	ttttgatctc	aacggatctc	tgctatccca	tgccacagaa	960
aaaatacacg	ctacgcggcg	cagcgggaag	cgcgtggaac	agtccagcca	ggagatctgg	1020
caggcggctc	gttcatgtat	tcgtaatgcg	ctcactctgg	cagacgtttg	tgacaaaagt	1080
gtggcaggca	tcggttttga	tgccacctgt	tctctggtgg	tactggataa	aaacggtgat	1140
ccattgcctg	tcagcccggg	aggagatgca	aagcaaaata	tcattgtgtg	gatggatcac	1200
cgcgccaccg	aacaagcaga	gcgaatcaat	gccactcacc	atccggtgct	gaactacgtc	1260
ggtggtaaaa	tttcgcctga	aatggaaaaa	ccgaaaaatt	cttggtcgaa	agaaaataatg	1320
ccagagatct	acgaacgtgc	cggacaattt	ttcgatctgg	cctgatttct	gacctggcgg	1380
gtaccgqgtg	atttaacgct	ttcagatgac	actgttacct	gtaaatggac	gtggctggca	1440

```

catgaaaatc gctgggatcc agattatttc cgcäaccattg gccttgacga gttagcggat 1500
gaagatttta ttcgcatggg tcatcatatt gtttctcccg gaacaccttg tggaaatggg 1560
ttaacagcac aagccgcggc agagatggga ttactccccg gcacacctgt cgtgttaggg 1620
ttaattgatg ctacgctggg tggcatcggt acggttggcg tagaagggtg agcgtgaac 1680
aatctcgctg atgttttcgg cacttcttca tgcaccatgg catctaccac ttctccctcg 1740
tttgtaccgg gtgtctgggg gccgtattac agtgcgatgg ttccaggctt gtggttagtt 1800
gaaggcgggc aaagtgtctg gggagcagct attgaccagc tacttgattt ccatccggct 1860
gttgaagaag ctgcgaaaat ggcacaacgt gtgaatcagc ccctccccgt ctggcttgct 1920
gatcgaatcc tcgaaaaaac ggcgcaacca tcagatgctg tcgccctggc gaaagggcta 1980
cacgtggtgc cggaatttct cggaatcgc gcccccttcg cagatcctca tgccagagcg 2040
gtaatttgtg gcctgggtat ggagcgagat ctggataatt tactcgctt gtatatcgct 2100
ggattatgcg gaattgggta tggctctgcg caaattctcg acgctcaaac agcgcagggg 2160
gtagtgagta aaaatatcgt tattagcggc ggtgcggggc agcatccact ggtacgacaa 2220
attctggcag atacctgctg tattccggtc attaccagc aatgctgcga accggtttta 2280
ttaggtctcg ccattcttgg tgcgtcgcc ggaaatattg caccttctgt tggcgaagcg 2340
atgcaacaat tcacccatgt ggataaatat tattatccgc aagaacgcta tcaatctctt 2400
catcatcgtc gatatgaggc ctataagcag ttgcagcata ctgcaaaatt actcagagac 2460
taattaacca gccgcctgac gctgttttca ggcaatcact aataacgact cactccggta 2520
atatcccgga gtgcattcat ctgcacccta aaaacgaggt ttatatgtcc agaaataata 2580
aacagtgggt gggtttgcca ctgcatctga tatggggata tatcgccatc gcagtattta 2640
tgactgggta tggtttcgaa ctgcattctt tatcgcata tattaatcg ttaggcttca 2700
cacctgcgga agcctctttt gcctttacgc tctacggcct ggcggtgctc ctttccgcct 2760
gggtttctgg ggtagtagcg gaaatcatca cgcgcacaaa aacctgctg attggttttg 2820
tcctatgggt cgtattccat gttctgtttc tggctcttgg attaggacag gcaaactatg 2880
ggttaatcct gctgttttac gggattcggt gtttagcgta tccgctatct ctctactcat 2940
ttatcgctgt tattattcat aacgtgcgta gcgaaaattc cagttctgct ctgggttggt 3000
actggcggtt atattcagta gggatcggtg ttgctggcag ttatattccc agttttacga 3060
taccgattat gggtgaaatg ggaaccttat ggctggcact ggcgttctgc tttgctggcg 3120
gtgtcattgc catgatctcc ttgcgtcatg ttaaaacgcc tggacatatg cataatttaa 3180
ctccccgtga gaagtttgca gaattaagtc gggcagtaac tttactttat accaaccgca 3240
atatttttct ctccagtatt gtgcgcatta taaatacctt atcgttattc ggttttgcg 3300
tcattatgcc aatgatgttt gtgatgaac tgggattcac cacctctgaa tggttgcagg 3360
tctggcgggc atttttcttt accactattt tctctaatat tttttggggg attgtggcag 3420
aaaaaatggg ctggatcgct gttattcgct ggtttgggtg cctcgggatg gcagcatcaa 3480
gtttagcggt ttactacatg ccgcaatact ttggtcacaa ctactggatg gcaatgattc 3540
cggcgattgc tctgggaact tttgttgctg catttgtgcc gatggccgct gtcttcccgg 3600
cactggaacc aaaacacaaa ggtgctgcaa tctcggttta caacctctct gcgggtatgt 3660
ctaacttcct ggctccggca attgccgtgg tgttattacc gtgggttagc actatcggtg 3720
tggtcattgc ctatacagca ttgtatctat tggcctttgt cctttgcgca ttcattcgcg 3780
ttgagcagcc aggattcagt tctgcgccag tgactgagaa ggcattgaat atctcctgaa 3840
aaacgaaacg catcaggcac tcatcctctc cctcatggga gaggatgatt tcacatcagg 3900
caatagtgc tttgttatcg agataaacgt cctgcacggc gttaatcagt ttcacgcgct 3960
cagccatcga t

```

<212> Type : DNA

<211> Length : 3971

SequenceName : SEQ ID NO 2

SequenceDescription : ribitol operon

Feature

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc_feature

<222> LocationFrom : 96

<222> LocationTo : 848

Other Information : ribitol dehydrogenase coding region

CDSJoin : No

Feature

Sequence: SEQ ID NO 2:

<221> FeatureKey : misc_feature

<222> LocationFrom : 859

<222> LocationTo : 2463

Other Information : ribitol kinase coding region

Feature

CDSJoin : No

Sequence Name : ribitol operon

SequenceDescription : ribitol dehydrogenase

SequenceDescription : ribitol kinase

SequenceDescription : ribitol transporter